種類(クラス数)の推定—パラメーターの巾和を用いる方法

志　村　利　雄

(1968年12月　受付)

A method of Estimation of Number of Classes

Toshio Shimura

We consider a multinomial distribution with infinitely many parameters $\{p_m\}$ and treat the problem of estimation of power sum such as $\sum_m' p_m{}^r$ or $\sum_{m_1 \neq m_2} p_{m_1}{}^{r_1} p_{m_2}{}^{r_2}$; the estinators are obtained very simply by means of a combinatorial method. Then we apply those estimators for a predictor of number of classes. As a corallary, we get simply the predictor obtained by Good and Toulmin [3], which is equivalent to our predictor in the sample of large size.

The Institute of Statistical Mathematics

1.　母集団 Ω が可附番無限個の部分集合（クラス）に分割されている，すなわち

$$\Omega = \bigcup_{m=1}^{\infty} \Omega_m \qquad \Omega_m \cap \Omega_{m'} = \phi \quad (m \neq m')$$

としよう．ここからランダムに大きさ n の独立な標本を複元抽出した観測値を $X_1, X_2, \cdots, X_n$ としたとき

$$P(X_j \in \Omega_m) = p_m \qquad m = 1, 2, \cdots;\ j = 1, 2, \cdots n$$
$$\sum_m' p_m = 1$$

となっているものとしよう．すなわち X_j は“可附番無限次元多項分布”に従う確率変数である．

標本の中にちょうど r 回出現したクラス Ω_m の数を n_r とし，出現したクラス数を d_n とすれば

$$n = \sum_{r=1}^{\infty} r n_r$$

$$d_n = \sum_{r=1}^{n} n_r$$

となる．

このような場合に取扱われている主な問題はつぎのようなものである (Harris [1])．

1)　大きさ αn $(\alpha \geq 1)$ の第2標本で観測されるクラス数 $d_{\alpha n}$ の予測．

2)　クラス出現範囲 (coverage) の推定．クラス出現範囲 C_n というのは

$$C_n = \sum_{m'}' p_{m'} .$$

ただし和は観測されたクラス $\Omega_{m'}$ の添字全体にわたる．

3)　大きさ αn の第2標本のクラス出現範囲 $C_{\alpha n}$ の予測．

このような問題について Harris [1], Good [2], Good and Toulmin [3], Goodman [4] にすぐれた結果が出ている．ここでは問題を母集団パラメーターの巾和との関係を取扱いながら主にクラス数にしぼって考察することにした．

2. 確率変数 $X_1, X_2, \cdots, X_n$ を用いれば,

$$d_n = \sum_m \left[1 - \prod_{j=1}^{n} (1 - I_{\Omega_m}(X_j))\right]$$

$$C_n = \sum_m p_m \left[1 - \prod_{j=1}^{n} (1 - I_{\Omega_m}(X_j))\right]$$

$$n_r = \sum_m I_{\{r\}}\left(\sum_{j=1}^{n} I_{\Omega_m}(X_j)\right) \qquad (r=1, 2, \cdots)$$

となる. ただし I_A は A の特性関数を示す. すなわち I_A は A 上で 1, A 以外で0である. これらの平均値は $X_1, X_2, \cdots, X_n$ の独立性から

$$E(d_n) = \sum_m (1 - (1 - p_m)^n)$$

$$E(C_n) = \sum_m p_m (1 - (1 - p_m)^n)$$

$$E(n_r) = \sum_m \binom{n}{r} p_m{}^r (1 - p_m)^{n-r}$$

となる.

まずわれわれは大きさ n の第1標本からの観測値 d_n, n_r, C_n をもとに大きさ $n\alpha$ $(\alpha \geq 1)$ なる第2標本から得られるはずの $d_{\alpha n}$, $C_{\alpha n}$ の値を予測しなければならない. d_n について考えるために $0 \leq t < \infty$ として関数

$$\varphi(t) = \sum_m \left[1 - (1 - p_m)^t\right]$$

を考察する. $\varphi(t) \geq 0$ かつ $\varphi(t)$ は単調増加である. 形式的に微分すれば

$$\varphi^{(r)}(t) = (-1)^{r-1} \sum_m (1 - p_m)^t \left(\log \frac{1}{1 - p_m}\right)^r.$$

$p_0 = \max p_m$ とおく ($p_0 < 1$ と仮定する). $\log(1-p_m)^{-1} \leqq p_m (1-p_m)^{-1} \leq p_0 (1-p_0)^{-1}$ だから

$$|\varphi^{(r)}(t)| \leq \frac{1}{(1 - p_0)^r}.$$

したがって $\varphi^{(r)}(t)$ は t に関して一様収束するから項別微分可能である.

$\varphi''(t) \leq 0$ であるから $\varphi(t)$ は $0 \leq t < \infty$ において上に凸である. したがって $t < s$ のとき

$$\frac{\varphi(s) - \varphi(t)}{s - t} \leq \varphi'(t).$$

ゆえに

$$\varphi(s) \leq \varphi(t) + (1 - t)\varphi'(t).$$

また

$$\varphi'(t) = \sum_m (1 - p)^t \left(\log \frac{1}{1 - p_m}\right) \leq \sum_m (1 - p_m)^{t-1} \cdot p_m$$

であるから, $t = n$, $s = \alpha n$ とおけば

(1) $$E(d_{\alpha n}) \leq E(d_n) + (\alpha - 1) n \sum_m (1 - p_m)^{n-1} p_m =$$

$$= E(d_n) + (\alpha - 1) E(n_1)$$

C_n については

$$E(C_n) = 1 - \sum_m p_m (1 - p_m)^n$$

$$= 1 - \frac{1}{n+1} E(n_1{}^{(n+1)}).$$

ただし, $n_1{}^{(n+1)}$ は大きさ $n+1$ の標本の中に1個だけ元が出現したクラス数である.

こういう論法をおしすすめて，$\varphi(t)$ を $t=n$ においてマクローリン級数に展開して $t=\alpha n$ とおけば

$$\varphi(\alpha n)=\varphi(n)+\sum_{r=1}^{\infty}\frac{(\alpha(n-1))!}{r!}(-1)^{r+1}\sum_{m}(1-p_m)^n\left(\log\frac{1}{1-p_m}\right)^r$$

となるから $d_{\alpha n}$ の展開式が得られるがこの各項には

$$\sum_{m}(1-p_m)^n\left(\log\frac{1}{1-p_m}\right)^r$$

という非常に推定困難なものがあらわれてくる．そのために，この展開を利用して問題解決することは得策でない．

3.　ここで観点を変えて推定を行うためにまずパラメーターの巾和について考察する．標本におけるクラス Ω_m の元の数を k_m とする，すなわち

$$k_m=\sum_{j=1}^{n}I_{\Omega_m}(X_j).$$

添字の集合 $\{1,2,\cdots,n\}$ から相異る r_1 個 $(1\leqq r_1<n)$ の添字 $(j_1,\cdots,j_{r_1})$ をえらび，残りの添字の集合から r_2 個の添字 $(j_{r_1+1},\cdots,j_{r_1+r_2})$ $(1\leqq r_2\leqq n-r_1)$ をえらぶことにすれば $\binom{n}{r_1}\binom{n-r_1}{r_2}$ 個の添字の組 $(j_1,\cdots,j_{r_1};j_{r_1+1},\cdots,j_{r_1+r_2})$ をえらぶことができる．このとき $m_1\neq m_2$ として

(3)　$$K(m_1,m_2;r_1,r_2)=\sum_{(j_1,\cdots,j_{r_1};j_{r_1+1},\cdots,j_{r_1+r_2})}I_{\Omega_{m_1}}(X_{j_1})\cdots I_{\Omega_{m_1}}(X_{j_{r_1}})I_{\Omega_{m_2}}(X_{j_{r_1+1}})\cdots I_{\Omega_{m_2}}(X_{j_{r_1+r_2}})$$

とおく，和は上述の $\binom{n}{r_1}\binom{n-r_1}{r_2}$ 個の添字の組 $(j_1,\cdots,j_{r_1};j_{r_1+1},\cdots,j_{r_1+r_2})$ 全体にわたる．

定　理　1　標本の大きさが n，標本の中に l 個の元が出現したクラス数が n_l のとき，自然数 r_1，r_2 $(1\leqq r_1,\ r_2<n,\ r_1+r_2\leqq n)$ に対して

(4)　$$P(r_1,r_2)=\binom{n}{r_1}^{-1}\binom{n-r_1}{r_2}^{-1}\sum_{l_1\geqq r_1}n_{l_1}\binom{l_1}{r_1}\left[\sum_{l_2\geqq r_2}n_{l_2}\binom{l_2}{r_2}-\binom{l_1}{r_2}\right]$$

は母集団パラメーターの巾の積和 $\sum_{m_1\neq m_2}p_{m_2}{}^{r_1}p_{m_1}{}^{r_2}$ の不偏推定値である．

証　明　$K(m_1,m_2;r_1,r_2)$ の右辺の和の各項で0でないものは $X_{j_1}\in\Omega_{m_1},\cdots,X_{j_{r_1}}\in\Omega_{m_1}$；$X_{j_{r_1+1}}\in\Omega_{m_2},\cdots,X_{j_{r_1+r_2}}\in\Omega_{m_2}$ なる添字の組 $(j_1,\cdots,j_{r_1};j_{r_1+1},\cdots,j_{r_1+r_2})$ をもつ項であるから $m_1\neq m_2$ に注意すれば $K(m_1,m_2;r_1,r_2)=\binom{k_{m_1}}{r_1}\binom{k_{m_2}}{r_2}$．他方 $K(m_1,m_2;r_1,r_2)$ の平均値をもとめれば (3) 式より

$$E\{K(m_1,m_2;r_1,r_2)\}=\binom{n}{r_1}\binom{n-r_1}{r_2}p_{m_1}{}^{r_1}p_{m_2}{}^{r_2}$$

となる．したがって，

$$P_1(r_1,r_2)=\binom{n}{r_1}^{-1}\binom{n-r_1}{r_2}^{-1}\sum_{m_1\neq m_2}K(m_1,m_2;r_1,r_2)$$

$$=\binom{n}{r_1}^{-1}\binom{n-r_1}{r_2}^{-1}\sum_{m_1\neq m_2}\binom{k_{m_1}}{r_1}\binom{k_{m_2}}{r_2}$$

とおけば

$$E\{P_1(r_1,r_2)\}=\sum_{m_1\neq m_2}p_{m_1}{}^{r_1}p_{m_2}{}^{r_2}.$$

すなわち $P_1(r_1,r_2)$ が $\sum_{m_1\neq m_2}p_{1m}{}^{r_1}p_{m_2}{}^{r_2}$ の不偏定値となる．$P_1(r_1,r_2)$ を書直せば

$$P_1(r_1, r_2) = \sum_{m_1 \neq m_2} \binom{k_{m_1}}{r_1}\binom{k_{m_2}}{r_2} = \sum_{m_1} \binom{k_{m_1}}{r_1} \sum_{m_2 \neq m_1} \binom{k_{m_2}}{r_2}$$

$$= \sum_{m_1} \binom{k_{m_1}}{r_1}\left[\sum_{m_2} \binom{k_{m_2}}{r_2} - \binom{k_{m_1}}{r_2}\right]$$

$$= \sum_{l_1 \geq r_1} n_{l_1} \binom{l_1}{r_1}\left[\sum_{l_2 \geq r_2} n_{l_2} \binom{l_2}{r_2} - \binom{l_1}{r_2}\right]$$

$$= P(r_1, r_2) \qquad \text{(証明終り)}$$

定 理 2 (Good [2]) n および n_l は定理1と同一，$r(1 \leq r \leq n)$ を自然数とすれば

$$P(r) = \binom{n}{r}^{-1} \sum_{k \geq r} n_k \binom{k}{r} \tag{5}$$

はパラメーターの巾和 $\sum_m p_m{}^r$ の不偏推定値である．

証 明 定理1と同様に証明できる．

つぎにパラメーターの巾和 $P(r)$ の分散をもとめておく．

定 理 3 n を標本の大きさとすれば (5) 式で定義した $P(r)$ の分散は

$$V(P(r)) = \frac{\binom{n-r}{r} - \binom{n}{r}}{\binom{n}{r}}\left(\sum_m p_m{}^r\right)^2 + \sum_{s=1}^{r} \frac{\binom{n}{s}\binom{n-s}{r-s}\binom{n-r}{r-s}}{\binom{n}{r}^2} \sum_m p_m{}^{2r-s} \tag{6}$$

となる．

証 明 添字の集合を $J=(j_1, \cdots, j_r)$ $I=(i_1, \cdots, i_r)$ とおき

$$X_m(J) = I_{\Omega_m}(X_{j_1}) \cdots I_{\Omega_m}(X_{j_r})$$
$$X_m(I) = I_{\Omega_m}(X_{i_1}) \cdots I_{\Omega_m}(X_{i_r})$$

とおけば $(K(m, r))^2$ を構成する和の項は $X_m(J)\ X_m(I)$ という形をしている．場合をわけて考える．

1) $n \leq 2r$ のときには $I \cap J = \phi$ なる項数は $\binom{n}{r}\binom{n-r}{r}$ 個ある．また $I \cap J \neq \phi$ のとき $\#(I \cap J) = s \leqq r$

($\#(\cdot)$ は $(\cdot)$ の中の集合の元の個数) なる項数は $\binom{n}{s}\binom{n-s}{r-s}\binom{n-r}{r-s}$ $(s = 1, 2, \cdots, r-1)$ 個ある．したがって

$$E(K(m, r)^2) = \binom{n}{r}\binom{n-r}{r} p_m{}^{2r} + \sum_{s=1}^{r} \binom{n}{s}\binom{n-s}{r-s}\binom{n-r}{r-s} p_m{}^{2r-s} \tag{7}$$

$$= \sum_{s=0}^{r} \binom{n}{s}\binom{n-s}{r-s}\binom{n-r}{r-s} p_m{}^{2r-s}$$

を得る．

2) $n < 2r$ のときにも二項係数の性質に注意すれば (5) 式の成立つことがわかる．

つぎに，$m_1 \neq m_2$ のとき $K(m_1, r)\ K(m_2, r)$ の平均値を求めなければならない．

$$K(m_1, r)\ K(m_2, r) = \sum_{I,J} X_{m_1}(I)\ X_{m_2}(J)$$

となるが，$m_1 \neq m_2$ に注意して $X_{m_1}(J)\ X_{m_2}(I) = 1$ となるためには，$I \cap J = \phi$ かつ $\#(I) = \#(J) = r$ でなければならない．したがって $m_1 \neq m_2$ のとき

$$E(K(m_1, r)\ K(m_1, r)) = \binom{n}{r}\binom{n-r}{r} p_{m_1}{}^r p_{m_2}{}^r$$

となる．

これだけの準備をしておいて $K(r)=\sum_m K(m,r)$ とおけば

$$K(r)^2=\sum_m K(m,r)^2+\sum_{m_1\neq m_2} K(m_1,r)\,K(m_2,r)$$

となるから

$$E(K(r)^2)=\sum_{s=0}^{r}\binom{n}{s}\binom{n-s}{r-s}\binom{n-r}{r-s}\sum_m p_m{}^{2r-s}+\binom{n}{r}\binom{n-r}{r}\sum_{m_1\neq m_2} p_{m_1}{}^r p_{m_2}{}^r.$$

また，$P(r)=K(r)\Big/\binom{n}{r}$ かつ

$$E(P(r)^2)=\sum_m p_m{}^{2r}+\sum_{m\neq m'} p_m{}^r p_{m'}{}^r$$

であるから

$$V(P(r))=\frac{\binom{n-r}{r}-\binom{n}{r}}{\binom{n}{r}}\left(\sum_m p_m{}^r\right)^2+\sum_{s=1}^{r}\frac{\binom{n}{s}\binom{n-s}{r-s}\binom{n-r}{r-s}}{\binom{n}{r}^2}\sum_m p_m{}^{2r-s}$$

を得る．（証明終り）

定理3と全く同様にして $P(r_1)$ と $P(r_2)$ の共分散をもとめることができる．

定　理　4　$P(r_1)$ と $P(r_2)$ の共分散は

(8)
$$\operatorname{cov}(P(r_1),P(r_2))=\frac{\binom{n-r_1}{r_2}-\binom{n}{r_2}}{\binom{n}{r_2}}\left(\sum_m p_m{}^{r_1}\right)\left(\sum_m p_m{}^{r_2}\right)+\sum_{s=1}^{r_1\wedge r_2}\frac{\binom{n}{s}\binom{n-s}{r_1-s}\binom{n-r_1}{r_2-s}}{\binom{n}{r_1}\binom{n}{r_2}}\sum_m p_m{}^{r_1+r_2-s}$$

となる．ここで $r_1\wedge r_2$ は r_1 と r_2 の最小値である．

さて，$E(d_{\alpha n})=\sum_m(1-(1-p_m)^{\alpha n})$ であったからこれを二項展開して

$$E(d_{\alpha n})=\sum_m\sum_{r=1}^{\alpha n}\binom{\alpha n}{r}(-1)^{r+1}p_m{}^r$$

$$=\sum_{r=1}^{\alpha n}(-1)^{r+1}\binom{\alpha n}{r}\sum_m p_m{}^r.$$

大きさ n の（第1）標本から作った $P(r)$（(5) 式参照）を $P_n(r)$ とし

(9)
$$\hat{d}_{\alpha n}=\sum_{r=1}^{n}(-1)^{r+1}\binom{\alpha n}{r}P_n(r)$$

を定義すれば

$$E(\hat{d}_{\alpha n})=\sum_{r=1}^{n}(-1)^{r+1}\binom{\alpha n}{r}E(P_n(r))=\sum_{r=1}^{n}(-1)^{r+1}\binom{\alpha n}{r}\sum_m p_m{}^r.$$

もし $E(d_{\alpha n})$ の predictor として (9) 式の $\hat{d}_{\alpha n}$ を使ったときその誤差すなわち $\hat{d}_{\alpha n}-E(d_{\alpha n})$ の二乗平均は

$$E[(\hat{d}_{\alpha n}-E(d_{\alpha n}))^2]=\sum_{r=1}^{m}\binom{\alpha n}{r}^2V(P_n(r))$$

$$+\sum_{r_1+r_2}(-1)^{r_1+r_2}\binom{\alpha n}{r_1}\binom{\alpha n}{r_2}\operatorname{cov}\left(P_n(r_1),\ P_n(r_2)\right)+\left(E(d_{\alpha n})-E(d_{\alpha n})\right)^2.$$

ここで

$$E(\hat{d}_{\alpha n})-E(d_{\alpha n})=\sum_{r=n+1}^{\alpha n}(-1)^{r+1}\binom{\alpha n}{r}\sum_m p_m{}^r.$$

4. (9) 式の $\hat{d}_{n\alpha}$ を $E(d_{n\alpha})$ の perdictor と考えたが n が大きいとき $\hat{d}_{n\alpha}$ はどうなるかしらべてみよう. いま r_0 は n に比較して小さくて, $P(r)=0$ $(r>r_0)$ であったとしよう. そうすれば n が大きいとき

$$\begin{aligned}
\hat{d}_{n\alpha}&=\sum_{r_0\geq r\geq 1}(-1)^{r+1}\binom{n\alpha}{r}P(r)\\
&=\sum_{r_0\geq r\geq 1}(-1)^{r+1}\binom{\alpha n}{r}\sum_{k\geq r}\frac{n_k\binom{k}{r}}{\binom{n}{r}}\\
&=\sum_{k\geq 1}n_k\left[\sum_{1\leq r\leq k}(-1)^{r+1}\frac{\binom{\alpha n}{r}\binom{k}{r}}{\binom{n}{r}}\right]\\
&\approx\sum_{k\geq 1}n_k\left[\sum_{1\leq r\leq k}(-1)^{r+1}\alpha^r\frac{1-\frac{r(r-1)}{2\alpha n}}{1-\frac{r(r-1)}{2n}}\cdot\binom{k}{r}\right]\\
&\approx\sum_{k\geq 1}n_k\sum_{1\leq r\leq k}(-1)^{r+1}\alpha^r\binom{k}{r}\\
&=\sum_{k\geq 1}n_k(-1)\left[(1-\alpha)^k-1\right]=d_n+\sum_{k\geq 1}(-1)^{k+1}(\alpha-1)^k n_k
\end{aligned}$$

よって $d_{n\alpha}$ の predictor として

$$\hat{\hat{d}}_{n\alpha}=d_n+\sum_{k\geq 1}(-1)^{k+1}(\alpha-1)^k n_k \tag{10}$$

も考えることができる. これは Good and Toulmin [3] で得られている predictor である.

われわれは $d_{n\alpha}$ の predictor を求めるためにもっぱら $\rho_r=\sum_m p_m{}^r$ $(r\geq 2)$ の不偏推定を用いるようにつとめてきた. しかし ρ_r は $r_0=\max\{k;\ n_k\neq 0\}$ とするとき $r\leq r_0$ についてのみ求められることができるが $r>r_0$ なときにも ρ_r を不偏推定することを考えなければならない. $\rho_r(r\leq r_0)$ の不偏推定を用うる限り (10) 式からもわかるように α が大きいと r_0 が偶数であるか奇数であるかによって推定が過大になったり過少になったりしてしまう. このような現象を防ぐためには, なるべく大きい r に対して ρ_r を推定することが望ましい. そのためにたとえば等式

$$\sum_m p_m{}^{r_0+1}=\sum_m p_m{}^{r_0}-\sum_{m_1\neq m_2}p_{m_1}{}^{r_0}p_{m_2}$$

を利用して ρ_{r_0+1} を推定することもできるであろう. すなわち標本を二分して一方から ρ_{r_0} を $P(r_0)$ で推定し, もう一方から $\sum_{m_1\neq m_2}p_{m_1}{}^{r_0}p_{m_2}$ を $P(r_0,1)$ で推定しその差 $P(r_0)-P(r_0,1)$ で ρ_{r_0+1} を推定する. こういう考え方でもっと大きい $r>r_0$ について ρ_r と推定することできよう. この場合には解決しなければならない問題点がいくつかある.

5. 簡単な数値例 (Harris [1] p. 541)

100個の相等しい出現確率

$$p_m=100^{-1}\ (m=1,2,\cdots,100)$$

をもった多項分布から大きさ 100 の標本をランダムに抽出したところつぎのようになった．

l	n_l
1	41
2	19
3	7
	$d_{100}=67$

このとき，

$P(1)=1$

$P(2)=0{,}00808081$

$P(3)=0{,}00004329$

$P(r)=0\ (r\geq 4)$

（実際には $\sum_m p_m{}^2=0{,}01$, $\sum_m p_m{}^3=0.0001$）

(1) 式による $d_{\alpha n}$ の上界 $\bar{d}_{\alpha n}$ と，(9) 式による $d_{\alpha n}$

$$\bar{d}_{\alpha n}=d_n+(\alpha-1)\,n_1$$

$$d_{\alpha n}=\sum_{r\geq 1}(-1)^{r+1}\binom{\alpha\, n}{r}P_n(r)$$

α	$\bar{d}_{\alpha n}$	$d_{\alpha n}$
2	108	97
3	148	130
4	190	213
5	231	388

この例では $d_{\alpha n}$ は $r_0=3$ までしか $P_n(r)\neq 0$ でないから α の最高次数が偶数である．そのためにすでに $\alpha\geqq 4$ で $\bar{d}_{\alpha n}\leqq d_{\alpha n}$ となってしまう．

統計数理研究所

文　献

[1] B. Harris, "Determining bounds on integrals with applications to cataloging problems" Ann. Math. Statist. Vol. 30 (1959)

[2] I. J. Good, "The population frequencies of spacies and the estimation of population parameters" Bionetrika, vol. 40 (1953)

[3] I. J. Good and G. H. Toulmin, "The number of new spacies and the increase in population coverage when a sample is increased" Biometrika, vol. 43 (1956)

[4] L. A. Goodman, "On the estimation of the number of classes in a population" Ann. Math. Stat. Vol. 20 (1949)

正　誤　表

彙報 15 巻 2 号 (1967) 訂正

"Composition and Rejection method" 法に関する注意

頁	行	誤	正
135	↑ 5	μ^* を F^*	μ^* を G
〃	↑ 2	前　節	上　記
136	↓ 15	$f(x)=\sum_{i=0}^{\infty}a_i f_i(x)$	$f(x)=\sum_{i=1}^{\infty}a_i f_i(x)\,g_i(x)$
〃	↑ 10	いま $F_i(x)$,	いま $F_0(x)$,
137	↑ 16	$N_N\leqq a$	$X_N\leqq a$
〃	↑ 10	一様乱数をとり	一様乱数 ξ をとり
138	↓ 9	$1\Big/\sum_{i=1}^{\infty}a_i$	$1\Big/\sum_{i=1}^{\infty}a_i$
〃	↑ 14	$f(x)=a_1f_1(x)+a_2f_2(x)\,g_2(x)$	$f(x)=a_1f_1(x)\,g_1(x)+a_2f_2(x)\,g_2(x)$